DØ note 4528-CONF

Measurement of the $t\bar{t}$ Production Cross-section at $\sqrt{s} = 1.96$ TeV in the $e\mu$ Channel Using Secondary Vertex b-tagging.

The DØ Collaboration
URL http://www-d0.fnal.gov
(Dated: August 3, 2004)

A measurement of the $t\bar{t}$ production cross-section at $\sqrt{s} = 1.96$ TeV in the $t\bar{t} \to e\mu$+jets channel using secondary vertex b-tagging is presented. The measurement is based on a 158 pb^{-1} data sample collected between June 2002 and September 2003. The preliminary cross-section obtained in this channel is:

$$\sigma_{t\bar{t}} = 11.1^{+5.8}_{-4.3} \text{ (stat)} \pm 1.4 \text{ (syst)} \pm 0.7 \text{ (lumi) pb.}$$

Preliminary Results for Summer 2004 Conferences

I. INTRODUCTION

The top quark is the heaviest of the known fundamental fermions. Measurements of its production and properties will therefore allow precision tests to be made of the theory of the strong interaction, quantum chromodynamics (QCD), and may provide sensitivity to new physics. This paper describes a measurement by the DØ experiment of the $t\bar{t}$ pair production cross-section in high energy $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV at the Tevatron collider.

At the Tevatron, the dominant production mechanism for top quarks is $t\bar{t}$ pair production, which occur via $q\bar{q}$ annihilation (85%) or gluon fusion (15%). Calculations of the $t\bar{t}$ cross-section at $\sqrt{s} = 2$ TeV using QCD next-to-leading order (NLO) range from 6.7 to 7.5 pb [1–3], if a top mass of 175 GeV is used. At next-to-next-to-leading order (NNLO) the predicted cross-section is 6.77 ± 0.42 pb [3].

In the Standard Model, the top quark decays almost with 100% probability to a W-boson and a b-quark. Therefore, in $t\bar{t}$ events the final state is completely determined by the decays of the two W-bosons. Experimentally, the most relevant channels are the "dilepton" channels where both of the W-bosons decay leptonically into an electron or a muon (ee, $\mu\mu$, $e\mu$) and the corresponding neutrino and the "lepton+jets" channels where one of the W-bosons decays hadronically into two jets (e+jets, μ+jets).[1] In all channels, two additional jets arise from the fragmentation of the b-quarks.

In order to measure the $t\bar{t}$ production cross-section, dedicated selections designed to enhance the signal content have to be applied. In the case of the $e\mu$ channel, the final state is characterized by a high p_T isolated electron, a high p_T isolated muon, large missing transverse energy and two b-jets. The presence of both an electron and a muon makes the signal very clean. The dominant physics backgrounds are $Z \to \tau\tau$ and WW. The dominant background owing to the imperfect reconstruction of the final state with the DØ detector (the so called instrumental background) is due to QCD multijet and W+jets production mechanisms in which one or both of the leptons are misidentified.

The signal for top production can be enhanced by exploiting differences in the heavy flavor content of the events containing top quarks compared with background events. This paper describes an analysis in which, for the first time, a b-tagging algorithm is used in the $e\mu$ channel to reconstruct displaced vertices.

II. THE DØ DETECTOR

The DØ Run II detector is comprised of the following main components: the central tracking system, the liquid-argon/uranium calorimeter, and the muon spectrometer.

The central tracking system includes a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located in a 2 T superconducting solenoid magnet. The SMT is designed to provide efficient tracking and vertexing capability at pseudorapidities of $|\eta| < 3$. The system has a six-barrel longitudinal structure, each with a set of four layers arranged axially around the beampipe, and interspersed with 16 radial disks. A typical pitch of 50-80 μ of the SMT strips allows a precision determination of the three-dimensional track impact parameter with respect to the primary vertex which is the key component of the lifetime based b-jet tagging algorithms. The CFT has eight coaxial barrels, each supporting two doublets of overlapping scintillating fibers of 0.835 mm diameter, one doublet being parallel to the collision axis, and the other alternating by $\pm 3^o$ relative to the axis [4].

The calorimeter is divided into a central section (CC) providing coverage out to $|\eta| \approx 1$, and two end calorimeters (EC) extending coverage to $|\eta| \approx 4$ all housed in separate cryostats. Scintillators placed between the CC and EC provide sampling of showers at $1.1 < |\eta| < 1.4$. [5]

The muon system, covering pseudorapidities of $|\eta| < 2$, resides beyond the calorimetry, and consists of three layers of tracking detectors and scintillating trigger counters. Moving radially outwards, the first layer is placed before the 1.8 T toroid magnets, and the two following layers are located after the magnets. [5]

III. MONTE CARLO MODELS FOR SIGNAL AND BACKGROUND SAMPLES

The $t\bar{t}$ signal expectations are determined from a full Monte Carlo simulation of top-antitop events. This simulation utilizes events generated at $\sqrt{s} = 1.96$ TeV with the ALPGEN 1.2 [6] matrix element generator assuming a top mass of 175 GeV/c^2 and the parton distribution function set CTEQ 6.1M [7]. These events are processed through PYTHIA 6.2 [8] to provide higher order QCD evolution (i.e. gluon radiation and fragmentation) and short lived particle decays.

[1] Unless otherwise explicitly stated a reference to a particle species can also be taken to refer to the anti-particle.

EvtGen [9] is used to model the decays of b hadrons. The W-bosons are both decayed to a lepton-neutrino pair, including all τ final states. The generated events are processed through a full detector simulation providing tracking hits, calorimeter cell energy and muon hit information. Multiple interactions are added to all events subject to Poisson statistics given the instantaneous luminosities typically observed in the data runs. The same reconstruction is applied to data and Monte Carlo events.

Although the main backgrounds are to a large extent studied using data, considerable resources are devoted to event simulations. The main physics background in the $e\mu$ channel, $(Z/\gamma* \rightarrow \tau\tau)jj$ is generated using ALPGEN followed by PYTHIA. To evaluate the Z background for lower jet multiplicities, PYTHIA samples are used. The other physics background considered, WW, is studied using $(WW \rightarrow ll)jj$ and $WW \rightarrow ll$ samples produced with ALPGEN followed by PYTHIA. In all above cases, τ's are forced to decay leptonically. In case of $Z/\gamma* \rightarrow \tau\tau$ the production cross-section measured by DØ (uncorrected for photon exchange and photon-Z interference) [10] is used. In case of $WW \rightarrow ll$, the next-to-leading-order cross-section (NLO) calculation [11] is used, instead of the leading-order ALPGEN cross-section, which represents an increase of 35% with respect to LO. For consistency, in the case of $(WW \rightarrow ll)jj$ the LO ALPGEN cross-section is scaled up by $(35 \pm 35\%)$ since no such NLO calculation is available for that process. For the $(Z/\gamma* \rightarrow \tau\tau)jj$ background, the ALPGEN cross-section, corrected for the difference in the yield for Z events observed in data and Monte Carlo, is used.

When utilizing simulated samples for analysis, additional smearing of the momenta of muons, electrons and jets is performed according to the observed resolution in data. Correction factors to efficiencies obtained on Monte Carlo are also applied, accounting for differences in, for instance, lepton identification efficiencies in data and Monte Carlo.

IV. THE DATA SET

The data sample used for this measurement consists of data collected during the period between June 2002 and September 2003. Only events from runs in which the tracking, calorimeter and muon systems were operating well are used. The integrated luminosity of the data set is 158 pb^{-1}.

V. OBJECT IDENTIFICATION

Electrons are identified as clusters in the electromagnetic layers of the calorimeter found within $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.4$ using a cone algorithm. These clusters are further required to be isolated from nearby hadronic energy, to satisfy a loose shower shape selection, and to satisfy a loose match to a central track. In the final selection a multiparameter likelihood discriminant is used which compares the probabilities for a certain candidate to be either a real electron or background (i.e. a highly electromagnetic jet). Electrons measured in the central ($|\eta| < 1.1$) and the forward ($1.5 < |\eta| < 2.5$) regions are used in this measurement. Energy scale corrections are derived using electrons reconstructed from Z boson decays.

Muons are built from a track segment in the inner muon layer matching a segment formed from hits in the outer two layers. The track formed in the muon system also has to be matched to a central track, and the overall χ^2 per number of degrees of freedom of the track fit is required to be less than 4. Timing cuts to reject cosmic ray muons are applied based on muon scintillator signals. Muons originating from decays of W or Z bosons are identified using two isolation criteria, one calorimeter based and one tracking based.

Jets are reconstructed using an iterative algorithm integrating energy observed in the calorimeter in a cone with radius $R = 0.5$. These jets are required to be in the region $|\eta| < 2.5$. The energy of the jets after reconstruction is corrected to represent the true jet energy. An equivalent correction is derived for Monte Carlo events. A further correction for heavy flavor jets is applied when a soft muon is found in a jet. This correction compensates for all of the muon energy as well as all of the corresponding neutrino energy.

The total missing transverse energy $\not\!\!E_{\rm T}$, is evaluated in order to provide sensitivity to the production of neutrinos with substantial p_T in an event. The change in each electron or jet p_T due to response corrections as described above is vectorially added to the $\not\!\!E_{\rm T}$, as are any observed muons.

The reconstuction of the primary vertex in the event, which is crucial for b-tagging, is made from central tracks which satisfy certain quality criteria for example on the number of hits in the SMT. Events are rejected if this vertex is not found within a longitudinal distance of less than 60 cm from the nominal interaction point, or if less than 3 tracks are used in its determination.

The DØ trigger is a three-level trigger system, where level 1 is a hardware trigger, while level 2 and 3 are software filters. The $e\mu$ trigger requires both an electron and a muon at the first trigger level, and one electron at the third trigger level. The trigger efficiency for the muon part is measured by analyzing events with calorimeter based triggers,

and for the electron part from events with muon based triggers. The trigger efficiencies for selected signal events with exactly one jet and two or more jets are $92.9 \pm 0.9\%$ and $92.8 \pm 0.5\%$ respectivelly.

VI. PRESELECTION OF $t\bar{t}$ EVENTS

The final state in the $e\mu$ channel is characterized by a high p_T isolated electron, a high p_T isolated muon, large missing transverse energy and two b-jets. The selection used is electron $p_T > 15$ GeV, muon $p_T > 15$ GeV, jet $p_T > 20$ GeV and $\not\!\!E_{\rm T} > 25$ GeV, where all the requirements on the electrons, muons and jets discussed in Section V are applied. The muon and electron tracks are further required to point to the primary vertex. Only events with one or more jets are considered.

VII. THE SECONDARY VERTEX TAGGING ALGORITHM

$t\bar{t}$ events contain two b-jets while jets in the background processes originate predominantly from light quarks or gluons. Requiring at least one jet in the event to be b-tagged is therefore a very powerful discriminant between signal and background. The b-tagging algorithm utilized in this measurement is the secondary vertex tagging algorithm (SVT), which explicitly reconstructs vertices that are displaced from the primary vertex. In the fitting procedure, the SVT only uses good quality tracks that have an impact parameter significance > 3. Only those fitted secondary vertices that are displaced from the primary vertex in the plane transverse to the beam line by more than 7 standard deviations are considered in this measurement. To take into account differences in the tracking performance in data and Monte Carlo, only jets containing at least two good quality tracks, with requirements for example on the number of hits in the SMT and the track p_T, are considered by the SVT. The fraction of these so called "taggable" jets, referred to as "taggability", is parameterized on data as a function of jet p_T and η, and applied to Monte Carlo.

The performance of the algorithm has been extensively studied in data. The semileptonic b-tagging efficiency is measured in a data sample enhanced in heavy flavor jets by requiring at least one jet in each event to contain a muon. This efficiency is then corrected to the inclusive efficiency using Monte Carlo. The c-tagging efficiency cannot be directly measured in data. Hence, it is obtained from the semileptonic b-tagging efficiency in data, corrected by the ratio of inclusive c-tagging efficiency to semileptonic b-tagging efficiency in Monte Carlo. The probability to tag a light flavor jet, the so called mistag rate, is also measured in data, from events where the secondary vertex has a negative decay length, meaning that the tracks forming the secondary vertex meet behind the primary vertex with respect to the jet axis. Both the b-tagging, c-tagging and light tagging efficiencies are parameterized as a function of jet p_T and η. These parameterizations are then applied to the Monte Carlo to predict the probability for a jet of a certain flavor to be tagged.

VIII. PHYSICS BACKGROUNDS

Background processes that can produce the full $e\mu$ signature (one electron, one muon, jets and significant missing transverse energy) are rare. Decays of $Z/\gamma*$ to two τ's which subsequently decay to an electron and a muon is the largest physics background, but suffers from the low branching ratio of the two τ's to decay to leptons, as well as from soft lepton and neutrino spectra. Diboson production, that of WW being the most important since the leptons resemble very much the ones in $t\bar{t}$ events, suffers from a very low cross-section. The most powerful discriminant in this measurement is however the requirement of at least one b-tagged jet.

The efficiencies for the two physics backgrounds considered to pass the preselection cuts (as defined in Section VI) are estimated from the Monte Carlo samples described in Section III. For the $Z/\gamma* \to \tau\tau$ background, the probability for such a background event to be b-tagged is estimated in data, using $Z \to ee$ and $Z \to \mu\mu$ events. For the small WW background the tagging efficiency is taken from Monte Carlo, by taking the jet kinematics from the Monte Carlo, and folding in the per jet tagging efficiency parameterizations determined on data (see Section VII). The tagging efficiencies for the physics backgrounds are listed in Table I.

IX. INSTRUMENTAL BACKGROUNDS

Additional background processes, for example from QCD multijet, W+jets and $Z \to \mu\mu$ production, can arise from misidentification of either the electron or the muon (or both). Electrons can be mimicked by photons or jets that fragment to a leading π^0 which then can acquire a track match due to overlap with a nearby charged hadron or from

photon conversion in the tracking volume. Non-isolated muons from jet fragmentation products or quark decay can occasionally appear isolated.

Fake lepton backgrounds are estimated by measuring the rate at which an electromagnetic jet is misidentified as an electron. This is done in a data sample where the muon in the event is non-isolated. The electron fake rate is then applied to the data sample that fulfills all requirements except that the electron identification cuts have been omitted. The efficiency for the very small $Z \to \mu\mu$ background (where one of the muons emits a Bremsstrahlung photon) to pass the preselection cuts is estimated using Monte Carlo.

The probability for a QCD or W+jets event to be b-tagged is estimated in data, from events containing an isolated muon passing all the muon identification cuts, and have $\not\!\!E_{\rm T} < 10$ GeV. The b-tagging efficiency for $Z/\gamma* \to \mu\mu$ background is determined in the same way as for the $Z/\gamma* \to \tau\tau$ background (as described Section VIII). The tagging efficiencies for the instrumental backgrounds are listed in Table I.

X. THE SIGNAL ACCEPTANCE

The preselection efficiency for $t\bar{t}$ events is estimated using the Monte Carlo sample described in III. The same Monte Carlo sample is used to estimate the tagging efficiency, by taking the jet kinematics from the Monte Carlo, and folding in the per jet tagging efficiency parameterizations determined on data (see Section VII). The tagging efficiencies for the $t\bar{t}$ signal are listed in Table I.

XI. SAMPLE COMPOSITION BEFORE AND AFTER b-TAGGING

Table I summarizes the per event tagging efficiencies for both the $t\bar{t}$ signal and the backgrounds, per jet multiplicity bin. Shown in Table II is the summary of the number of expected and observed events both before and after b-tagging. The expectation for $t\bar{t}$ is calculated assuming a production cross-section of 7 pb. The summary of the observation as well as the number of predicted signal and background events are shown in Fig. 1 as a function of jet multiplicity.

	Tagging efficiency (%)		
Sample	Njets = 0	Njets = 1	Njets $\geq$ 2
$t\bar{t}$		38.0 $\pm$ 2.1	58.9 $\pm$ 1.6
WW		0.34 $\pm$ 0.06	1.8 $\pm$ 0.3
$Z \to \tau\tau$		0.8 $\pm$ 0.2	2.9 $\pm$ 1.0
$Z \to \mu\mu$		0.8 $\pm$ 0.2	2.9 $\pm$ 1.0
QCD and W+jets		1.6 $\pm$ 0.1	3.3 $\pm$ 0.5

TABLE I: The per event tagging efficiencies for the $t\bar{t}$ signal as well as all the backgrounds. Only statistical errors are shown.

XII. $t\bar{t}$ PRODUCTION CROSS-SECTION

The $t\bar{t}$ production cross-section is calculated by maximizing the product of the likelihood functions constructed for each of the two jet multiplicity bins (Njets = 1 and Njets $\geq$ 2). The likelihood function is based on the Poisson probability to observe a given number of events given the expected number of signal and background events. The resulting likelihood as a function of the $t\bar{t}$ production cross-section is shown in Fig. 2.

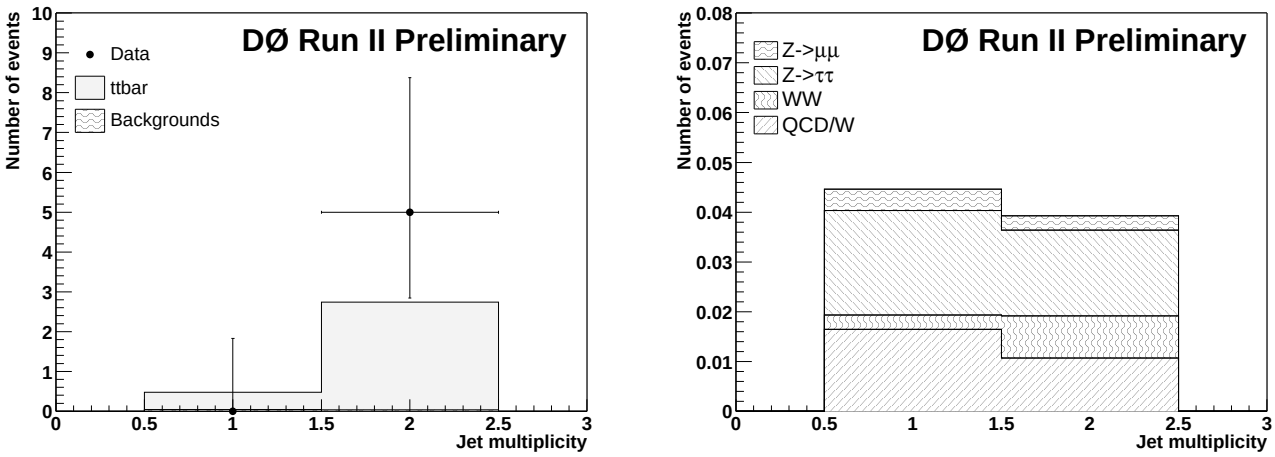
The systematic uncertainty on the measured cross-section is obtained by varying the background prediction and the signal efficiencies, within their errors, with all the correlations between the two jet multiplicity bins as well as between the different classes of backgrounds taken into account. Uncertainties from limited Monte Carlo and data statistics are treated as uncorrelated. Table III summarizes all the main systematic uncertainties. The systematic uncertainity on the luminosity measurement is 6.5%.

The preliminary cross-section measured in the $e\mu$ channel is:

$$\sigma_{t\bar{t}} = 11.1^{+5.8}_{-4.3}\ (\text{stat}) \pm 1.4\ (\text{syst}) \pm 0.7\ (\text{lumi})\ \text{pb}.$$

Sample	Njets = 0	Njets = 1	Njets $\geq$ 2
	Expected number of events before tagging		
$t\bar{t}$	0.052 $\pm$ 0.009	1.14 $\pm$ 0.04	4.58 $\pm$ 0.09
WW	7.40 $\pm$ 0.08	0.84 $\pm$ 0.03	0.46 $\pm$ 0.03
$Z \to \tau\tau$	1.6 $\pm$ 0.1	2.7 $\pm$ 0.2	0.6 $\pm$ 0.1
$Z \to \mu\mu$	1.2 $\pm$ 0.4	0.6 $\pm$ 0.3	0.10 $\pm$ 0.04
QCD and W+jets	2.9 $\pm$ 0.1	1.05 $\pm$ 0.07	0.33 $\pm$ 0.04
Total	13.2 $\pm$ 0.4	6.3 $\pm$ 0.3	6.1 $\pm$ 0.2
	Observed number of events before tagging		
	13	7	8
	Expected number of tagged events		
$t\bar{t}$		0.43 $\pm$ 0.03	2.70 $\pm$ 0.09
WW		< 0.005	0.008 $\pm$ 0.002
$Z \to \tau\tau$		0.021 $\pm$ 0.006	0.017 $\pm$ 0.007
$Z \to \mu\mu$		< 0.005	< 0.005
QCD and W+jets		0.016 $\pm$ 0.002	0.011 $\pm$ 0.002
Total		0.48 $\pm$ 0.03	2.74 $\pm$ 0.09
	Observed number of tagged events		
		0	5

TABLE II: The expectation and observation before and after b-tagging as a function of jet multiplicity. Only statistical errors are shown.



(a) The observation together with the signal and background predictions.

(b) The background prediction.

FIG. 1: The number of observed and predicted b-tagged events (left) and the predicted number of background events (right), as a function of jet multiplicity.

XIII. SUMMARY

Using the $t\bar{t} \to e\mu$+jets channel, the cross-section for $t\bar{t}$ pair production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV has been measured by the DØ experiment to be $11.1^{+5.8}_{-4.3}$ (stat) ± 1.4 (syst) ± 0.7 (lumi) pb. The data agrees with perturbative QCD calculations.

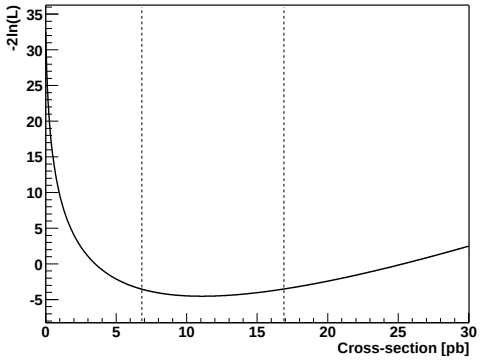


FIG. 2: The likelihood function obtained using both the 1- and 2-jet bins, plotted versus the cross-section. The vertical lines indicate the statistical error on the measured cross-section.

Systematic	Δ cross-section (pb)
b-tagging efficiency in data	+0.75 -0.64
b-tagging efficiency in Monte Carlo	+0.48 -0.46
Decay model dependence of tagging efficiency	-0.32
Taggability	+0.39 -0.80
Jet energy scale	+0.51 -0.38
Jet energy resolution	-0.023
Jet identification	+0.26
Trigger	+0.34 -0.27
Top mass	+0.43 -0.37
Monte Carlo to data correction factors	$\pm$ 0.46
Monte Carlo and data statistics	$\pm$ 0.34
Total	$\pm$ 1.4

TABLE III: Summary of systematic uncertainties in pb.

Acknowledgments

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the Department of Energy and National Science Foundation (USA), Commissariat à L'Energie Atomique and CNRS/Institut National de Physique Nucléaire et de Physique des Particules (France), Ministry for Science and Technology and Ministry for Atomic Energy (Russia), CAPES, CNPq and FAPERJ (Brazil), Departments of Atomic Energy and Science and Education (India), Colciencias (Colombia), CONACyT (Mexico), Ministry of Education and KOSEF (Korea), CONICET and UBACyT (Argentina), The Foundation for Fundamental Research on Matter (The Netherlands), PPARC (United Kingdom), Ministry of Education (Czech Republic), Natural Sciences and Engineering Research Council and WestGrid Project (Canada), BMBF (Germany), A.P. Sloan Foundation, Civilian Research and Development Foundation, Research Corporation, Texas Advanced Research Program, and the Alexander von Humboldt Foundation.

[1] E. Berger and H. Contopanagos, Phys. Rev. D **57**, 253 (1998).
[2] R. Bonciani, S. Catani, M. Mangano and P. Nason, Nucl. Phys. B529, 424 (1998).
[3] N. Kidonakis and R. Vogt, Phys. Rev. D **68**, 114014 (2003).
[4] V. Abazov et al. (DØ Collaboration), in preparation for submission to Nucl. Instrum. Methods Phys. Res. A; T. LeCompte and H.T. Diehl, "The CDF and DØ Upgrades for RunII", Ann. Rev. Nucl. Part. Sci. **50**, 71 (2000)
[5] S. Abachi et al. (DØ Collaboration), Nucl. Instrum. Methods Phys. Res. A **338**, 185 (1994).
[6] M.L. Mangano et al., JHEP 0307:001, 2003, hep-ph/0206293.
[7] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. Nadolsky, W. K. Tung, hep-ph/0201195.
[8] T. Sjöstrand et al., LU TP 01-21, hep-ph/0108264.
[9] http://charm.physics.ucsb.edu/people/lange/EvtGen/

[10] E. Nurse and P. Telford, DØ note 4284.
[11] J.M. Campbell and R.K. Ellis, Phys Rev D **60**, 113006 (1999).

APPENDIX A: KINEMATIC DISTRIBUTIONS

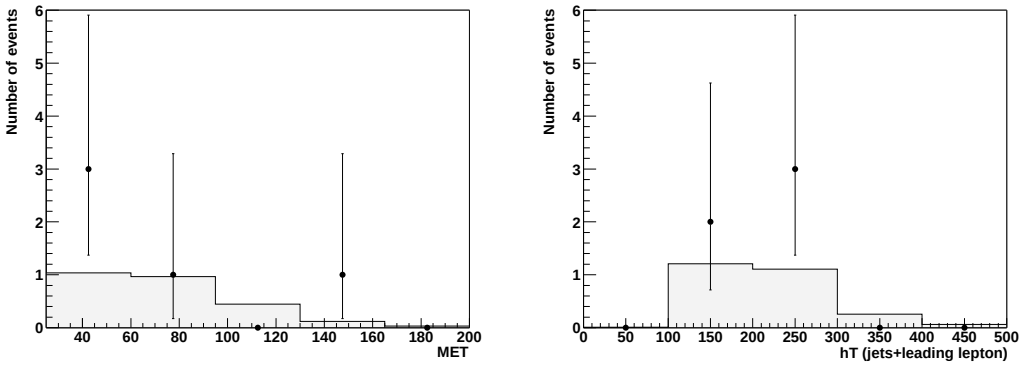


FIG. 3: Number of observed and predicted signal events in the second jet multiplicity bin, as a function of $\not\!\!E_{\rm T}$ and H_T (the prediction is shown for the $t\bar{t}$ signal only).

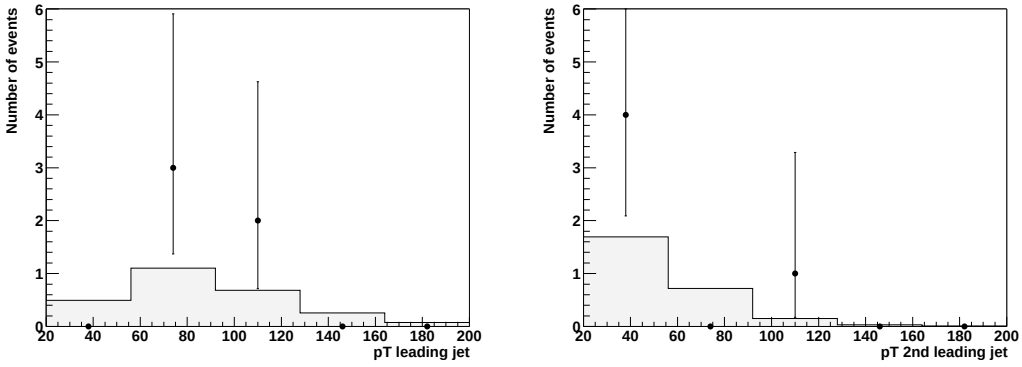


FIG. 4: Number of observed and predicted signal events in the second jet multiplicity bin, as a function of p_T of the leading jet and second leading jet (the prediction is shown for the $t\bar{t}$ signal only).

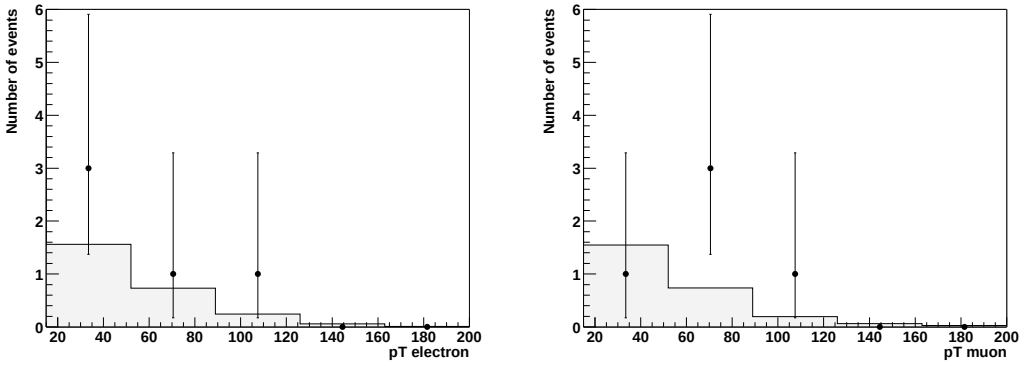


FIG. 5: Number of observed and predicted signal events in the second jet multiplicity bin, as a function of p_T of the electron and the muon (the prediction is shown for the $t\bar{t}$ signal only).

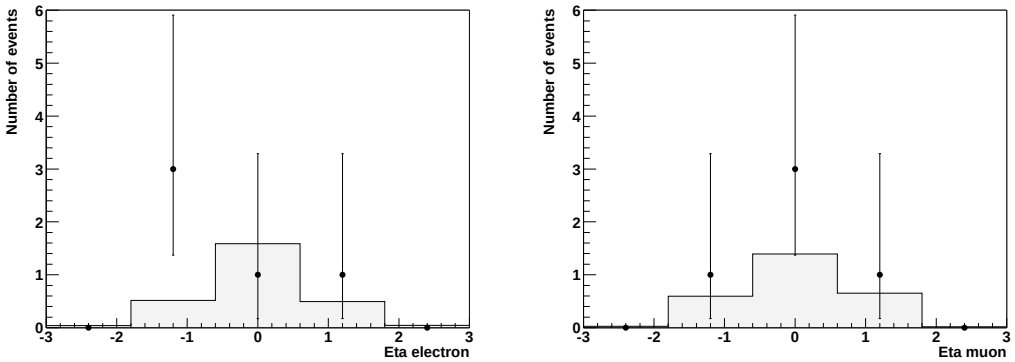


FIG. 6: Number of observed and predicted signal events in the second jet multiplicity bin, as a function of η of the electron and the muon (the prediction is shown for the $t\bar{t}$ signal only).